

**PCT**WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau

## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |           |                                                                                                                                                                                                                                                                                                                                                                                                |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>(51) International Patent Classification <sup>6</sup>:</b><br><b>C12N 15/11, 15/12</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | <b>A1</b> | <b>(11) International Publication Number:</b> <b>WO 97/08307</b><br><b>(43) International Publication Date:</b> 6 March 1997 (06.03.97)                                                                                                                                                                                                                                                        |
| <b>(21) International Application Number:</b> PCT/KR96/00145<br><b>(22) International Filing Date:</b> 23 August 1996 (23.08.96)<br><b>(30) Priority Data:</b><br>1995/26391 24 August 1995 (24.08.95) KR<br><b>(71) Applicant (for all designated States except US):</b> IL DONG PHARMACEUTICAL CO., LTD. [KR/KR]; 60, Yangjae-dong, Seocho-ku, Seoul 137-130 (KR).<br><b>(72) Inventors; and</b><br><b>(75) Inventors/Applicants (for US only):</b> KIM, Sun, Young [KR/KR]; 3-306, Hanshin Seolae Apartment, 70-1, Panpo-dong, Seocho-ku, Seoul 137-040 (KR). KIM, Kee, Won [KR/KR]; 77-1402, Hyundai Apartment, 456, Apkujjeong-dong, Kangnam-ku, Seoul 135-110 (KR). KIM, Tae, Han [KR/KR]; 206-1503, LG Apartment, 111, Jeongja-dong, Puntang-ku, Seongnam-shi, Kyonggi-do 463-010 (KR). HWANG, Jeong, Ho [KR/KR]; 203-1202, Hanshin Apartment, 124, Imae-dong, Puntang-ku, Seongnam-shi, Kyonggi-do 463-060 (KR). KIM, Seon, Hee [KR/KR]; 23-704, Hanshin 3rd Apartment, Panpo 2-dong, Seocho-ku, Seoul 137-042 (KR). LEE, Sun, Young [KR/KR]; 205, Samick Hights Villa, 38-3, Karak-dong, Songpa-ku, Seoul 138-160 (KR). |           | <b>(74) Agent:</b> KIM, Won, Ho; You Me Patent & Law Firm, 702, Teheran Boulevard, 825-33, Yoksam-dong, Kangnam-ku, Seoul 135-080 (KR).<br><b>(81) Designated States:</b> AU, CA, CN, JP, KR, MX, RU, US, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).<br><b>Published</b><br><i>With international search report.</i><br><i>With amended claims.</i> |
| <b>(54) Title:</b> HETEROLOGOUS PROTEIN PRODUCTION SYSTEM USING AVIAN CELLS                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |           |                                                                                                                                                                                                                                                                                                                                                                                                |
| <b>(57) Abstract</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |           |                                                                                                                                                                                                                                                                                                                                                                                                |
| <p>Heterologous protein expression system including a heterologous gene DNA such as EPO genomic DNA, a vector receiving the DNA and an avian cell, such as duck embryo or quail fibrosarcoma cell line, expressing the gene in the vector can be used to efficiently produce heterologous proteins such as EPO.</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |           |                                                                                                                                                                                                                                                                                                                                                                                                |

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

|    |                          |    |                                       |    |                          |
|----|--------------------------|----|---------------------------------------|----|--------------------------|
| AM | Armenia                  | GB | United Kingdom                        | MW | Malawi                   |
| AT | Austria                  | GE | Georgia                               | MX | Mexico                   |
| AU | Australia                | GN | Guinea                                | NE | Niger                    |
| BB | Barbados                 | GR | Greece                                | NL | Netherlands              |
| BE | Belgium                  | HU | Hungary                               | NO | Norway                   |
| BF | Burkina Faso             | IE | Ireland                               | NZ | New Zealand              |
| BG | Bulgaria                 | IT | Italy                                 | PL | Poland                   |
| BJ | Benin                    | JP | Japan                                 | PT | Portugal                 |
| BR | Brazil                   | KE | Kenya                                 | RO | Romania                  |
| BY | Belarus                  | KG | Kyrgyzstan                            | RU | Russian Federation       |
| CA | Canada                   | KP | Democratic People's Republic of Korea | SD | Sudan                    |
| CF | Central African Republic | KR | Republic of Korea                     | SE | Sweden                   |
| CG | Congo                    | KZ | Kazakhstan                            | SG | Singapore                |
| CH | Switzerland              | LI | Liechtenstein                         | SI | Slovenia                 |
| CI | Côte d'Ivoire            | LK | Sri Lanka                             | SK | Slovakia                 |
| CM | Cameroon                 | LR | Liberia                               | SN | Senegal                  |
| CN | China                    | LT | Lithuania                             | SZ | Swaziland                |
| CS | Czechoslovakia           | LU | Luxembourg                            | TD | Chad                     |
| CZ | Czech Republic           | LV | Latvia                                | TG | Togo                     |
| DE | Germany                  | MC | Monaco                                | TJ | Tajikistan               |
| DK | Denmark                  | MD | Republic of Moldova                   | TT | Trinidad and Tobago      |
| EE | Estonia                  | MG | Madagascar                            | UA | Ukraine                  |
| ES | Spain                    | ML | Mali                                  | UG | Uganda                   |
| FI | Finland                  | MN | Mongolia                              | US | United States of America |
| FR | France                   | MR | Mauritania                            | UZ | Uzbekistan               |
| GA | Gabon                    |    |                                       | VN | Viet Nam                 |

## HETEROLOGOUS PROTEIN PRODUCTION SYSTEM USING AVIAN CELLS

### BACKGROUND OF THE INVENTION

#### 5 1. Field of the Invention

The present invention relates to novel expression systems that can produce biomedically important heterologous proteins including human erythropoietin (hereafter "EPO"), and more specifically to the production of various heterologous proteins by transfecting DNA  
10 encoding the proteins, such as the genomic DNA encoding EPO into avian cells.

#### 2. Related Arts

Many recombinant proteins used in medicine are relatively small and simple in their structure, and biologically functional proteins can be  
15 produced in prokaryote such as *E. coli*. However, some human proteins of medical interest, such as TPA (tissue plasminogen activator), Factor VIII, EPO, etc. are more complicated because biological function requires post-translational modification. For example, EPO is extensively glycosylated with the carbohydrate portion  
20 accounting for 40 % of the molecular mass. It has been shown that the carbohydrate portion of EPO is important for biological function. Accordingly, EPO produced in *E. coli*, yeast or insect is inactive or very weakly active *in vivo*, while EPO produced in COS or CHO cells was found to be fully active. Accordingly, those kinds of heterologous  
25 proteins have been produced only in mammalian cells.

In the meantime, the avian system has been used for the study of gene expression in higher eukaryote for a long time. One of the first viruses to be linked to tumors was the *Rous sarcoma* virus of chicken,

and this virus was instrumental in demonstrating that the retroviral oncogene can originate from a cellular gene, leading to the concept of the protooncogen. Studies of gene expression have also been done using the RSV LTR promoter, which has often be used for high level  
5 expression of heterologous genes in mammalian cells. In addition, avian embryo cells have been used extensively in studies of various animal viruses.

### SUMMARY OF THE INVENTION

The present invention is a research for the high level expression  
10 of eukaryotic heterologous proteins. It is an object of the present invention to provide a novel heterologous gene expression system which can produce proteins of higher eukaryotic cells. It is another object to provide the method of efficiently producing higher eukaryotic proteins, such as EPO, etc., which has been known to be active only  
15 when they are produced in a mammalian cell. It is a further object of the invention to provide the method of producing, especially, EPO among the eukaryotic proteins described above.

To accomplish the objects of the present invention, the present invention provides a heterologous gene expression system comprising  
20 a DNA encoding a heterologous protein, a vector for receiving the DNA; and an avian cell for harboring the vector.

The present invention also provides a method of producing a heterologous protein comprising the steps of culturing the expression system of claim 1 in media to express the heterologous gene, and  
25 purifying the heterologous proteins from the cell and the media.

Preferably, the heterologous protein of the present invention is selected from the group consisting of those proteins that are known to be active only when expressed in mammalian cells (such as EPO, TPA,

Factor VIII, etc.) and preferably, the vector contains a promoter selected from the group consisting of SV early promoter, major immediate early promoter of human cytomegalovirus (hereafter "HCMV MIEP") and RSV LTR, and preferably, the avian cell is selected from the group consisting of duck embryo cell (hereafter "DE"), chicken embryo fibroblast (hereafter "CEF") and quail fibrosarcoma (hereafter "QT"), more preferably QT-VC which was isolated by the inventors. QT-VC was deposited to the International Depository Authority, Korea Research Institute of Bioscience and Biotechnology Korean Collection for Type Culture, and assigned a deposit number of KCTC 0277BP on August 22, 1996. The deposited QT-VC was transfected with the expression vector containing SY-EPO cDNA as described in Fig. 8.

More preferably, the DNA encoding the heterologous protein is genomic DNA or cDNA.

Further, the present invention provides an EPO production system comprising a DNA encoding EPO, a vector for receiving the DNA, and an avian cell for harboring the vector.

Moreover, the invention provides a method of producing EPO comprising the steps of inserting a DNA encoding EPO into a vector, transfecting the vector into an avian cell, and culturing the transfected avian cell in media.

Preferably, the avian cell of the EPO production system is DE or QT, and the DNA is a genomic DNA encoding EPO, more preferably, the DNA selected from the group consisting of SY, JM, SH and HE described in Fig. 5.

Preferably, the vector has a promoter selected from the group consisting of SV early promoter, HCMV MIEP and RSV LTR.

The present invention also provides an avian cell as a host for

expressing genes encoding mammalian proteins.

Further, the present invention provides an novel EPO genomic sequence selected from the group consisting of SY, JM, SH and HE described in Fig. 5, and also provides an novel EPO amino acid  
5 sequence selected from the group consisting of JM, SH and HE described in Fig. 6.

#### BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 shows the expression of the bacterial CAT gene in avian cells. DE and CEF cells were transfected with pRc/CMV containing  
10 (+) or lacking (-) the CAT sequence. CAT activity was measured by determining the amount of acetylated chloramphenicol (AC) produced from  $^{14}\text{C}$ -chloramphenicol. The values shown are from one representative of more than five independent assays. For this particular experiment, 10  $\mu\text{g}$  of protein was reacted with  $^{14}\text{C}$ -  
15 chloramphenicol for 20 min at 37 °C.

Fig. 2 shows the comparison of CAT gene expression between various cell types and between different promoters. The three promoter-CAT fusion constructs were transfected into DE, CEF, CHO-K1, and HeLa cells, and CAT activity was measured as described in  
20 Fig. 1. S, SV40 early promoter; C, HCMV MIEP; R, RSV LTR. The values shown are from one representative of three independent assays. For this particular experiment, 10  $\mu\text{g}$  of protein was reacted with  $^{14}\text{C}$ -chloramphenicol for 30 min at 37 °C .

Fig. 3 shows the efficiency of DNA transfection in various cells.  
25 pCMV-lacZ constructs was transfected into DE, CHO, Vero, HeLa, and 293T cells by calcium phosphate-DNA coprecipitation using the conditions used for the experiments shown in Fig. 2. Two days after transfection, cells were fixed and stained with X-gal. The number of

blue cells per 60 mm tissue culture plate was counted. The total number of cells between plates were comparable at  $1-3 \times 10^5$ . Transfection efficiency was calculated relative to DE cells.

Fig. 4 shows the schematic diagram for cloning of human EPO and construction of expression vectors. The five blocks represent the five coding regions of EPO. The first PCR was performed using primers 25 and 33. The amplified DNA fragment was cloned and subjected to a second PCR using primers 12 and 9. The wavy tale in primer 12 contains the nucleotide sequence from the first coding region. Therefore, the second PCR generates the entire coding sequence of EPO so that the first and the second coding regions are attached to form without intron between them. Primers 12 and 9 contain HindIII linkers at their 5' ends, enabling cloning of the EPO genomic sequence into various expression vectors.

Fig. 5 is various EPO genomic DNA sequences. SY, SH, HE and JM are the EPO genomic DNA sequences cloned by the present invention, and AM and GI are the EPO genomic sequences which has been already reported. Since the intron between the first coding region and the second coding region was deleted during the cloning, the deleted intron is not shown in Fig. 5.

Fig. 6 is various EPO amino acid sequences. SY, SH, HE and JM are the EPO amino acid sequences cloned by the present invention, and AM and GI are the EPO amino acid sequences which have been already reported. The abbreviation of the amino acids are as follows:

A: alanine R: arginine N: asparagine D: aspartic acid  
C: cystein Q: glutamine E: glutamic acid H: histidine  
I: isoleucine L: leucine K: lysine M: methionine

F: phenylalanine P: proline S: serine

T: threonine W: tryptophan Y: tyrosine V: valine

Fig. 7 shows the comparison of CAT gene expression between QT-VC and other mammalian cell lines. pCMV-CAT was transfected to QT-VC; CHO-K1, and Vero cells, and CAT activity was measured as described in Fig. 1. The transfection efficiency, as measured by X-gal staining following cotransfection with pCMV-lacZ, was reproducibly 3-5 % in all cases. For this particular experiment, 50 µg of protein were incubated with <sup>14</sup>C-chloramphenicol for one hour at 37 °C.

Fig. 8 is the typical structure of a plasmid used to express EPO in QT cells. The two types of BamHI cassettes which could express the gene for human glutamine synthetase (GS) was made. In these BamHI cassettes, the GS cDNA sequence was flanked by the poly A sequence from the bovine growth hormone gene and one of the two promoters, the partial MMTV LTR (from -220 to +15 from the RNA start site) or the 220 bp HSV tk promoter. The BamHI fragment expressing GS was inserted into the BamHI site of pCI-neo (Promega, Madison, WI, USA), resulting in a series of pIGA. The HindIII fragment of the SY-EPO cDNA sequence was cloned into the SmaI site of pIGA, generating the EPO expression vector, pIGA-EPO.

Fig. 9 shows the production of EPO by QT-N4D4. QT-N4D4 cells were grown to confluence in a 10 cm culture dish (day 0) in M-199 containing 10 % FBS and 1 mM MSX. On day 3, the EPO level was measured. The cells were then split into 1:3 and seeded onto 10 cm dishes. On day 6, the cells were again reached confluence, and the medium was replaced with 10 ml fresh medium containing 2 % (●) or 10 % (○) FBS. EPO levels were determined by ELISA (R & D system, Minnesota, USA)

Fig. 10 shows the comparison of EPO concentration in DE (●) and QT-N4D4 (○) measured by ELISA and by *in vitro* bioassay.

#### DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

The inventors have explored the possibility of using avian cells as a host cell for heterologous gene expression. We have chosen to use three avian cells; two embryonic cells from chicken and duck, and a quail fibrosarcoma line. We chose to use the chicken and duck embryo cells for the following the reasons. First, these embryonic cells can easily be prepared from eggs, and they divide rapidly, undergoing many passages. Second, chicken and duck cells can be grown at large scale with relatively low costs. Third, some avian cells, such as those from chicken embryos have already been used for medical products. For example, influenza virus has been cultured in chicken eggs for the production of vaccines. Finally, the culture conditions, including media and temperature, required by avian embryo cells are virtually identical to those of mammalian cells, suggesting that the physiology of avian and mammalian cells is probably comparable. Further, the reason of choosing a QT cell line is that various transformed cell lines have been already developed, and it is easy to handle these cell lines to construct a permanent cell line expressing a heterologous protein, and the culture conditions and media is similar to those of mammalian cells.

#### I. Cells and Plasmids

##### 1. Cells

The following Table 1 shows cells used in the experiment.

Table 1

| Cells                                                                              | Source                                         |
|------------------------------------------------------------------------------------|------------------------------------------------|
| HeLa human cervical carcinoma cells                                                | ATCC CCL2                                      |
| Vero African green monkey kidney cells                                             | ATCC CCL81                                     |
| COS-7 African green monkey kidney cells transformed by wild-type T antigen of SV40 | ATCC CRL1651                                   |
| CHO-K1 Chinese hamster ovary cells                                                 | ATCC CCL61                                     |
| NIH3T3 contacted-inhibited Swiss mouse embryo cells                                | ATCC CRL1651                                   |
| Ad-5 transformed human embryonic kidney cells 293                                  | ATCC CRL1651                                   |
| SL-29 chicken embryo fibroblast cells                                              | ATCC CRL1590                                   |
| Duck embryo                                                                        | ATCC CCL141<br>or prepared by the<br>inventors |
| Quail fibrosarcoma line QT6                                                        | ATCC CRL1708                                   |
| Quail fibrosarcoma line QT-VC                                                      | Isolated by the inventors<br>KCTC 0277BP       |

All these cells except QT cell lines were grown in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum (FBS). QT cell lines were cultured in M199 medium instead of DME. Duck embryo was either obtained from ATCC CCL 141 or prepared by trypsinization of 10- to 13- day old decapitated duck embryos. These avian cells were grown in minimum essential medium (Eagle) supplemented with non-essential amino acids and Earle's balanced salt solution containing 10 % FBS. These cells could be

maintained for approximately another 30 passages. Each medium used in this study was supplemented with 120 µg/ml penicillin G (Sigma P-3032; 1690 units per mg) and 200 µg/ml streptomycin (Sigma S-9137; 750 units per mg).

5           2. Plasmids

To evaluate the efficiency of heterologous protein production in avian cells, pRc/RSV-CAT and pRc/CMV-CAT were constructed by inserting a HindIII CAT cassette (Pharmacia, Piscataway, NJ) into the HindIII sites of pRc/RSV and pRc/CMV (Invitrogen, San Diego, California, USA), respectively. For pSVCAT, the plasmid p918 was used, which has been already described by the inventors. For EPO expression vectors, three vectors were used. pCMV-gEPO was constructed by cloning the HindIII fragments of the EPO genomic sequence into the HindIII site of pRc/CMV. pSV-gEPO was derived by replacing the CAT sequence of pSV918 with the genomic EPO sequence. pIGA-EPO has cDNA of EPO controlled by HCMV MIEP and the genes of NEO and glutamine synthetase (hereafter "GS"). To measure the transfection efficiency, the plasmid pCMV-lacZ was constructed by inserting bacterial lacZ fragment into the HindIII site of pRc/CMV.

II. DNA Transfection and Gene Expression Assays

The inventors tested whether avian embryo cells could be used for high levels of heterologous gene expression instead of mammalian cells. Although avian embryo cells have been used to culture viruses, there was no report that heterologous proteins of higher eukaryotic cells were expressed in these cells. To carry out the study, it is necessary to develop the method of efficient transfection to avian cells. That is, to express heterologous genes in avian cells, it is required to develop

the transfection technique of DNA to target cells. At present, we could not find any reports on DNA transfection of avian embryo cells. Accordingly, the inventors have developed the technique that CEF and DE cells can readily be transfected with DNA.

5        Among the techniques available, we have chosen a method using calcium phosphate coprecipitation, because this works well for various adherent cells and can also be used for establishing permanent lines. We have tested many different conditions and found that the following procedure was optimum.

10        When cultures were 50-70% confluent in a 100 mm culture dish, a total of 10 µg DNA in HBS buffer (140 mM NaCl, 5 mM KCl, 0.75 mM  $\text{Na}_2\text{HPO}_4 \cdot 2\text{H}_2\text{O}$ , 6 mM dextrose, 25 mM HEPES) was incubated with the cells for 30 min at room temperature. 10 ml of regular media containing FBS was added and incubated for 20 hrs at 37 °C, except for  
15        CHO-K1 (8 hours). Cells were then treated with 10 ml of 100 µM chloroquine, and incubated for another 3 hours at 37 °C. After replacement with 10 ml of fresh media, the cells were grown for 1 to 2 days. Culture supernatants were collected and centrifuged at 1000 rpm for 10 min to remove cells and debris. To measure transfection  
20        efficiency, cells were transfected with pCMV-lacZ, rinsed once with PBS 3 days after transfection, fixed with 0.5 % glutaraldehyde (in PBS) for 10 min, and washed twice for 2-10 min each with 4 ml PBS containing 1 mM  $\text{MgCl}_2$ . For X-gal staining, the staining solution [PBS  
25        containing 4 mM  $\text{K}_3\text{Fe}(\text{CN})_6$ , 4 mM  $\text{K}_4\text{Fe}(\text{CN})_6 \cdot 3\text{H}_2\text{O}$ , 2 mM  $\text{MgCl}_2$ , and 400 µg per ml X-gal (in dimethylformamide)] was added to fixed cells, and incubated at 37 °C for 4 hours overnight. When the reaction was completed, cells were washed once with PBS. Stained cells were kept in PBS.

CAT assay was carried out as follows:

Two to three days after transfection, cells were harvested, washed once with PBS, and resuspended in 0.25 M Tris-HCl (pH 7.5).

Total proteins were prepared by 4 cycles of freeze/thawing followed by heating at 65 °C for 7 min. Equivalent amounts of protein were assayed for CAT activity at 37 °C for 30 min. The amount of protein and the reaction time varied, depending on the experiments. For example, the CAT activity of cell extracts prepared from DE cells was so high that only 10 µg protein and 20 to 30 min reaction time had to be used, and under this condition, levels of CAT activity in other mammalian cells were very low or undetectable. When CAT activity became detectable in other cells, virtually all <sup>14</sup>C-chloramphenicol was converted. The percent conversion of <sup>14</sup>C-chloramphenicol to its acetylated forms was determined by cutting out regions containing unreacted and acetylated forms and quantifying the amount of radioactivity in each by liquid scintillation counting.

### III. Gene Expression in DE and CEF

Gene expression efficiency of DE and CEF was measured using CAT gene. We initially chose to use a promoter from the major immediate-early region of HCMV, because this has been shown to drive a high level gene expression in many different cell types. In the plasmid pCMV-CAT, the bacterial CAT gene is placed under the control of the HCMV MIEP. As a negative control, the plasmid Rc/CMV containing the promoter but no CAT sequence was used. These plasmids were transfected into DE and CEF cells and CAT activity was measured to estimate the efficiency of transfection and gene expression. One representative result from several independent transfections is shown in Fig. 1. Transfection of a control plasmid resulted in undetectable levels of CAT activity in both cells. However,

transfection with pCMV-CAT resulted in readily detectable levels of CAT activity in both cells. In more than five independent transfection assays, the level of CAT activity was always higher in DE cells than in CEF cells. The magnitude of difference in the level of CAT activity  
5 between the two cells ranged from 10- to 50-fold, depending on the experiment. This result indicated that avian cells were readily transfected with DNA and the heterologous genes could be efficiently expressed.

#### 10 IV. Comparison of Levels of Gene Expression between Avian and Mammalian Cells, and between Different Promoters

We have compared the levels of gene expression between avian and mammalian cells, using three different promoters;

- (1) the SV40 early promoter, which is used during the early transcriptional phase of SV40 infection;
- 15 (2) the HCMV MIEP, which drives the expression of IE1 and IE2 regulatory proteins, immediately after HCMV infection;
- (3) the RSV LTR from an avian retrovirus.

These promoters are known to be powerful in mammalian cells, and have often been used for high level heterologous gene expression.

20 These promoter-CAT fusion constructs were transfected into four different cell lines, DE, CEF, CHO-K1, and HeLa, and CAT activity measured to compare the efficiency of gene expression between promoters and between cell types. To make this comparison semi-quantitative, all transfections and CAT assays were performed at  
25 the same time and using identical conditions. One representative result of such experiments is shown in Fig. 2. Here, 10 µg of cell extracts were incubated for 30 min in the CAT reaction. Under these

particular conditions, the levels of CAT expression driven from the three promoters were very low in CHO and HeLa cells (Fig. 2). Only after larger amounts of proteins were used for extended reaction time, was CAT activity detected. In contrast, CAT activity was readily detectable in the avian cells (Fig. 2), except for the SV40 promoter in CEF cells. It indicated that the expression in avian cells are more effective than that in mammalian cells.

The most dramatic finding was that the HCMV MIEP was extremely powerful in DE cells. In Fig. 2, the conditions used for the CAT reaction were chosen to generate the reasonable levels of CAT activity in other samples. When the CAT reaction was performed under limiting conditions for the protein sample prepared from DE cells transfected with pCMV-CAT (i.e., when the CAT conversion was below 50 %), the levels of CAT activity of all the other samples were virtually undetectable. Therefore, the magnitude of difference in CAT activity between the protein sample from DE cells transfected with HCMV-CAT and those from the other transfections is at least two orders of magnitude. These results suggested that heterologous genes could be expressed very efficiently under the control of the HCMV MIEP in DE cells.

It is possible that the high levels of CAT expression seen in DE cells could be due to efficient transfection of the cell population, rather than an ability of these cells to support strong gene expression. To distinguish these possibilities, we transfected pCMV-lacZ into DE and various animal cells. After transfection, cells were stained with X-gal, and the number of blue cells were counted to estimate the transfection efficiency. As shown in Fig. 3, the number of stained cell was always comparable between DE and other animal cells, suggesting that the high levels of CAT expression in DE cells were due to high levels of

expression in individual cells.

#### V. Cloning of human erythropoietin

To test whether DE cells could indeed be used for the expression of medically important human proteins, we have isolated the genomic DNA encoding the human EPO gene. We chose to use EPO  
5 as a model because it is a secreted protein, so we could test whether DE cells properly process secreted proteins. We also used a genomic clone of EPO instead of the cDNA, to assess whether human genes are properly spliced to produce functional mRNAs in DE cells.

10 DNAs for cloning of EPO were prepared with blood cells collected from four people. Human peripheral blood lymphocytes were isolated by Ficoll-Hypaque gradient centrifugation of heparin-treated blood cells. Total DNA was prepared and used for polymerase chain reaction using specific oligonucleotide primers (Fig.  
15 4). The region around the start codon was highly GC rich, so the EPO sequence was cloned by two steps of PCR using two different pairs of primers.

To obtain the genomic DNA for EPO, total DNA was prepared by lysing human peripheral blood lymphocytes using TES (10 mM Tris-HCl  
20 pH 7.8; 1 mM EDTA; 0.7 % SDS) followed by the treatment with 400 µg/ml proteinase K at 50 °C for 1 hour, phenol:chloroform extraction, and ethanol precipitation. The polymerase chain reaction (PCR) was performed using 0.1 µg of total genomic DNA and oligonucleotide primers specific to the EPO gene.

25 Primer #25 (sense, 5' to 3'): GAAGCTGATAAGCTGATAACC

Primer #33 (antisense, 5' to 3'): TGTGACATCCTTAGATCTCA

The samples were amplified through 30 cycles that included the

following parameters; denaturation at 92 °C for 1 min, primer annealing at 55 °C for 1min, and primer extension at 72 °C for 1 min. The DNA fragment amplified from this reaction did not contain the first 13 nucleotides in the N-terminal region, so a second PCR was performed  
5 using the following primers (Underlined, HindIII; Outlined, start codon and stop codons, respectively). The relative position of these primers are as shown in Fig. 4. Taq DNA polymerase (POSCO Chem, Korea) and pfr polymerase (STRATGENE, California, USA) were used to amplify DNA.

10 Primer #12 (sense, 5' to 3'):

CAAGCTTCGGAGATGGGGTGACGAATGTCCTGCCTGGCTGTGGC

Primer #9 (antisense, 5' to 3'):

CAAGCTTTCATCTGTCCCCTGTCCTGC

The amplified DNA from the second PCR was cloned into the  
15 pCRII (Invitrogen), from which the HindIII fragment containing the genomic sequence of EPO was inserted into various expression vectors as described above. In this experiment, the amplified DNA was placed under the control of the HCMV MIEP or SV40 early promoter, generating pCMV-gEPO and pSV-gEPO respectively. SY-  
20 EPO whose amino acid sequence is identical to that of the already known EPO is used for the expression experiments in the sections VII and VIII (See the section VI).

#### VI. Analysis of Nucleotide Sequences of Cloned EPO Genomes

Genomic structure of EPO cloned by the above method is different from  
25 the natural EPO genome *in vivo*. That is, wild type EPO genomic DNA has five coding regions and four introns between them. However, in the DNA cloned by the above method, the first coding region was fused

to the second coding region to form one coding region so that it has four coding regions and three introns (Fig. 4).

The results from the analysis of EPO gene sequences isolated from four people suggested that nucleotide sequences of EPO cloned from  
5 these region are significantly different from those of the prior two EPOs (AM-EPO and GI-EPO) (Fig. 5) at the sites of intron. Such a difference was not due to the error which occurred during DNA amplification in the process of cloning. We repeated cloning and sequencing using DNAs prepared from same individuals (but at  
10 different times) and obtained the same nucleotide sequence. As another control, we amplified the already cloned EPO under the similar conditions, and determined the nucleotide sequence. Again, we obtained the same nucleotide sequence.

Amino acid sequences of four EPO genes, together with AM and GI,  
15 are shown in Fig. 6. Amino acid sequences from AM, GI and SY are identical. However, amino acid sequences from three people (JM, SH, HE) different by two or three different amino acids from GI- and AM-EPO, suggesting that there is a polymorphisms among people. When compared with AM- or GI-EPO, HE-EPO had three different amino  
20 acids at C-terminal, SH-EPO three different amino acid over the whole polypeptide, and JM-EPO two different amino acids, one at C-terminal and the other in the middle of polypeptide (See Fig. 6). For example, while AM-EPO and GI-EPO had serine, alanine, and valine at positions 36, 100 and 170 respectively, SH-EPO had arginine, serine, and  
25 tyrosine. Further, while AM-EPO and GI-EPO had valine, lysine, and aliginine at positions 170, 177, and 191, HE-EPO had tyrosine, glutamine, and glycine. In JM-EPO, lysine and tyrosine were present at positions 54 and 170, while they were threonine and valine. These results suggested that the EPO gene has a polymorphism in amino

acids sequence as well as DNA sequence.

## VII. Expression of EPO in DE Cells

In this experiment, we compared levels of EPO expression between DE cells and other cell lines.

5 EPO expression vectors were transfected into various cells including DE, CEF, CHO, HeLa, VERO, and 293T. We have included VERO cells because they are often used for heterologous gene expression, and 293T cells which drive very high levels of gene expression, presumably due to both the high frequency of DNA  
10 transfection and the presence of potent viral transactivators such as EIA, EIB, and large T antigen. Two to three days after transfection, levels of EPO in the culture supernatants were measured by the enzyme linked immunoadsorbent assay, and transfection efficiencies were determined by staining cells adhered on the culture with X-gal.  
15 Transfection efficiency was carried out by transfection of a lacZ expression vector together with an EPO expression vector as described in the section II. One representative result of this analysis is summarized in Table 2.

Table 2

| Cell   | HCMV MIEP | SV40 early promoter | HCMV/SV40 |
|--------|-----------|---------------------|-----------|
| 293    | 314       | 17.5                | 18        |
| CHO    | 139.4     | 10.4                | 13.5      |
| VERO   | 250       | 10.7                | 23.5      |
| NIH3T3 | 89        | 79.4                | 1.1       |
| DE     | 4335      | 13.8                | 314.8     |

20

When the SV40 early promoter was used, there was little

difference in the levels of EPO between cell types. However, when the HCMV MIEP was used, DE cells produced much higher levels of EPO than any other cell lines tested. The HCMV MIEP was much more active than the SV40 early promoter in almost all the cells tested.

5 This difference was especially pronounced in DE cells, where the former produced 315 times more EPO than the latter. Among the various cell types, DE cells always produced the highest level of EPO.

CHO cells are the source of cell lines producing EPO that is currently used for human application. In this transient system, however, the  
10 level of EPO in CHO cells was at least 30-fold lower than in DE cells. Difference between DE and 293T cells was also considerable. Transfection efficiency of 293T was higher by about 30-fold than any other cells including DE cells. Moreover, 293T cells produce potent viral transcription transactivators. Nevertheless, DE produced 10-  
15 fold more EPO than 293T, suggesting that DE could drive high levels of the gene expression.

In conclusion, human EPO could efficiently be produced and secreted in DE cells and that the HCMV MIEP is the promoter of choice for driving high level heterologous gene expression in DE cells.

20 In summary, we found that DE cells could produce very high levels of bacterial and human proteins. All three promoters tested drove higher levels of gene expression in DE cells than any other cell lines used in this study. In particular, the HCMV MIEP was extremely powerful in DE cells. The high level of heterologous gene expression  
25 observed was not due to a higher number of transfected cells. It appears that DE cells properly process splicing and secretion because transfection of DE cells with an expression vector containing the EPO genomic DNA sequence produced a large quantity of EPO in the culture supernatant.

For DE cells to be used for industrial purpose, one would need to develop large-scale culture techniques for these cells. There are two possible ways. First, it may be possible to use primary cells themselves as the producer line. A large number of DE cells can easily be prepared from 10- to 13 day-old duck embryos. From one embryo, we can readily obtain  $10^9$  to  $10^{10}$  cells that can undergo at least 15 passages. Therefore, it is possible to transfect DE cells at the earliest possible stage with an expression vector followed by selection of transfected cells, which might require 4-7 passages. Even if less than 5% of the cells were transfected, a large number of transfected cells would be available, suggesting that large-scale culture of primary duck embryo cells is not impossible with primary cells. Second, it will be possible to transform duck embryo cells at an early stage, using one of the large number of well-characterized oncogenes that are available. With transformed DE cells, a producer line could be constructed, and better quality control of protein production be established. It remains to be seen whether transformed DE cells will still maintain the capability for high level gene expression. Although a number of biological questions remain to be answered, the potential of these cells for the production of various proteins warrants further investigation.

#### VIII. Heterologous Gene Expression in the Transformed Avian Cell Line

The above experiments demonstrated the great potential of DE cells as producers of heterologous proteins such as EPO. However, DE cells used in the above experiments are primary cells and stop dividing after 30-40 passages *in vitro*. Therefore, unless DE cells are transformed or special techniques are developed as described above, it is difficult to use these embryonic cells for industrial production of heterologous proteins.

In the following study, we tested whether the transformed avian cell line, namely the quail fibrosarcoma line, could be used to produce EPO. The quail fibrosarcoma line used in this study, QT-VC, was subcloned from QT6 (ATCC CRL1708). This line was derived from methylcholanthrene-induced fibrosarcoma of Japanese quail. QT-VC is different from its parental line in at least two aspects. First, QT-VC grows faster than the parental line in M199 medium containing 10% FBS used in this study. The former divided every 12-24 hours, while the doubling time of the latter was 24-36 hours. Second, the QT-VC cell looks more roundish than QT6 which generally grows in a longish form. Like its parental line, QT-VC did not grow well when it was seeded at a low density. Therefore, cells had to be split to 1/3 to 1/2 after reaching confluence for continuous culture.

#### 1. Analysis of Gene Expression in QT-VC Cells

We compared the levels of gene expression between QT-VC and mammalian cells using pCMV-CAT. We chose to use the HCMV MIEP as this promoter was shown to drive high levels of gene expression in various cell types including avian cells (See the section IV). pCMV-CAT was transfected into 3 cell lines, QT-VC, CHO-K1 and Vero. To make this comparison semi-quantitative, all transfections and CAT assays were performed at the same time and using identical conditions. Transfection efficiency was also measured by cotransfecting pCM-lacZ followed by X-gal staining. The efficiency was approximately 3 % in all cases. Under these conditions, the levels of CAT expression in QT-VC cells were always 2-3 times higher than mammalian cell lines used in this study (Fig. 7). Although the level of gene expression in QT-VC cells appears to be lower than DE cells, the quail fibrosarcoma line is at least as good as mammalian cell lines, suggesting that it could be used as a producer for heterologous

proteins.

## 2. Construction of EPO Expression Vectors for QT-VC Cells

To test whether high levels of heterologous proteins could be expressed in QT cells, we have constructed other EPO expression  
5 vectors. The basic strategy for the construction of an expression vector was as follows:

First, we chose to use the HCMV MIEP to drive expression of the heterologous gene as it had already been shown to be one of the strongest promoters in avian cells as well as mammalian cells.

10 Second, the human glutamine synthetase (GS) gene was used for amplification of the target gene. Generally, the gene of interest is amplified to augment the yield of protein by using certain selectable markers in the presence of specific chemicals. One of the best examples is the dihydrofolate reductase (DHFR) gene. It has been  
15 shown that the copy number of the heterologous gene and the level of respective protein increase as the concentration of methotrexate (MTX) in the medium is slowly increased. However, this system requires the host cell defective in the gene DHFR, so cannot be directly applied to QT cells for which such a mutant line is not yet available. For this  
20 reason, we chose to use the GS gene. In this case, the host cell line need not to be deficient for GS, because only multiple copies of the GS gene can confer resistance to methionine sulfoximine (MSX).

The overall structure of EPO expression vectors constructed for the use in QT cells is shown in Fig. 8. In this structure, the cDNA  
25 sequence for EPO is under the control of the HCMV MIEP, the bacterial Neo gene is used as the first selectable marker, and the human GS gene is also present as the second selectable marker in the same plasmid. The backbone of expression vectors used in this particular

experiment was pCI-neo (Promega, USA) which uses the HCMV MIEP and the intron from the  $\beta$ -globin genome. We have made a couple of different constructs in which the GS gene is driven by the partial MMTV LTR (from -220 to +15) or the 220 bp HSV tk promoters. In either  
5 case, the magnitude of gene amplification appears to be comparable (Data not shown). Detailed procedure and supplementary data regarding the construction of expression vectors is available upon request.

### 3. Construction of QT-VC Cells Stably Expressing EPO

10 To construct QT-VC cell lines constitutively expressing EPO, the cells were transfected with an EPO expression vector by a calcium phosphate coprecipitation method as described in the section II. Three days after transfection, EPO production was confirmed by EILSA and transfected cells were treated with G418 (0.8 mg/ml) and MSX (25  
15  $\mu$ M). When G418-resistant cells were grown to confluence, cells were diluted for subcloning. Because QT-VC cells do not grow efficiently at a low cell density, cells were seeded on 10-cm culture dishes at various numbers ( $10^2$ ,  $10^3$ ,  $10^4$ ,  $10^5$  per dish). Then the colonies that grew distant from other colonies were isolated by plastic O rings and  
20 expanded onto a 96-well plate. When cells reached 70 % confluence, the EPO level was measured. Subclones that produced more than 200 U/ml were serially expanded from the 12-well to 6-well to 60 mm culture plates. When cells reached confluence on a 60 mm dish, cells  
25 were split on 6-well plates and then treated with various concentrations of MSX (100  $\mu$ M, 250  $\mu$ M, 1 mM). Using this procedure, several subclones that produced large amounts of EPO and also grew fast were selected.

One of the subclones obtained through this procedure is QT-

N4D4. As shown in Fig. 9, this subclone produced 1200 U/ml when grown for 3 days after confluence. When the cells were split to 1:3, seeded on 10 cm dishes, and allowed to grow for another 3 days, N4D4 still produced 1000 U/ml. The medium was then replaced with a fresh media containing 2 % FBS and the cells still produced 400 U/ml EPO.

These results indicated that QT cells could produce a large quantity of EPO.

In conclusion, the above experiment demonstrated the great potential of QT cells as a producer for heterologous protein.

#### IX. Biological Activity of EPO Produced in Avian Cells

EPO is heavily glycosylated and such glycosylation is required for its biological activity. For example, EPO produced in *E. coli* or yeast is inactive or very weakly active *in vivo*. To test whether EPO expressed in DE or QT cells was biologically active, we carried out an *in vitro* bioassay using spleen cells isolated from mice treated with phenylhydrazine.

EPO assay: Absolute levels of EPO production after transfection of various cells were determined by enzyme linked immunoadsorbent assay which is currently used to measure EPO levels in the human serum (R & D Systems Inc., Minnesota, USA). To measure the biological activity of EPO, *in vitro* bioassay was carried out by the method of Krystal as modified by Goldberg et al Spleen cells were taken from C57BL X C3H FI hybrid mice (Seoul National University Laboratory Animal Center) on day 3 after the second of two daily injections of phenylhydrazine (60 mg/Kg of body weight per day) and spleen cell suspensions were prepared with Lymphoprep™ (NYCOMED PHARMA AS, Oslo, Norway). The spleen cells (final concentration  $4 \times 10^6$  cells per ml) were then incubated in 24 well tissue

culture plates with various standard doses of EPO (CILAG AG International, Switzerland; specific activity 2000 U/ml) or unknown samples for 22 hr and then pulsed with 4  $\mu$ Ci/well tritiated thymidine (Amersham Co.) for 2-3 hr. The cells were harvested, washed with  
5 PBS several times and lysed by 0.3 N NaOH and 0.1% SDS. Radioactivity in LSC cocktail solutions were calculated by a Pharmacia Wallac 1410 scintillation counter.

Culture supernatants from QT-N4D4 cells or DE cells transfected with EPO expression vectors were taken to measure levels of EPO by  
10 both ELISA and the bioassay. The ELISA measures absolute concentration, and is currently used for determining EPO concentration in human serum. On the other hand, the bioassay determines biological activity using a control EPO that has been produced from mammalian cells and is currently being used in humans. Fig. 10  
15 compares the difference in levels of EPO determined by these two methods. The ratio between the values (mU) was  $1 \pm 0.15$ , and the specific activity of EPO produced from DE cells was estimated to 105 U/ $\mu$ g. Therefore, the levels of EPO measured by ELISA were very comparable to those obtained by the bioassay. This result suggested  
20 that EPO produced from these avian cells had a similar biological activity to commercially available EPO.

What is claimed is:

1. A heterologous gene expression system comprising:  
a DNA encoding a heterologous protein;  
a vector for receiving the DNA; and  
5 an avian cell for harboring the vector.
2. The expression system of claim 1, wherein the heterologous protein is selected from the group consisting of TPA, Factor VIII and EPO.
3. The expression system of claim 1, wherein the vector  
10 contains a promoter selected from the group consisting of SV early promoter, HCMV MIEP and RSV LTR.
4. The expression system of claim 1, wherein the avian cell is selected from the group consisting of DE, CEF and QT.
5. The expression system of claim 4, wherein the QT is QT-VC.
- 15 6. The expression system of claim 1, wherein the DNA encoding the heterologous protein is DNA or cDNA.
7. An avian cell as a host for expressing a gene encoding a mammalian heterologous protein.
8. A method of producing a heterologous protein comprising  
20 the steps of;  
culturing the avian cell containing the expression system of claim 1 to express the gene of the heterologous protein in media; and  
purifying the heterologous protein from the cell and the media.
9. The method of claim 8, wherein the heterologous protein is  
25 selected from the group consisting of TPA, Factor VIII and EPO.

10. The method of claim 8, wherein the vector contains a promoter selected from the group consisting of SV early promoter, HCMV MIEP and RSV LTR.

11. The method of claim 8, wherein the avian cell is selected  
5 from the group consisting of DE, CEF and QT.

12. The method of claim 11, wherein the QT is QT-VC.

13. The method of claim 8, wherein the DNA encoding the heterologous protein is DNA or cDNA.

14. An EPO production system comprising:  
10 a DNA encoding EPO;  
a vector for receiving the DNA; and  
an avian cell for harboring the vector.

15. The EPO production system of claim 14, wherein the avian cell is DE or QT.

16. The EPO production system of claim 15, wherein the QT is  
15 QT-VC.

17. The EPO production system of claim 14, wherein the DNA is a genomic DNA encoding EPO.

18. The EPO production system of claim 14, wherein the DNA  
20 encoding EPO is selected from the group consisting of SY, JM, SH and HE described in Fig. 5.

19. The production system of claim 14, wherein the vector contains a promoter selected from the group consisting of SV early promoter, HCMV MIEP and RSV LTR.

20. A method of producing EPO comprising the steps of:  
25

inserting a DNA encoding an EPO into a vector;  
transfecting the vector into an avian cell; and  
culturing the transfected avian cell in media.

21. The method of claim 20, wherein the avian cell is DE or QT.

5 22. The method of claim 21, wherein the QT is QT-VC.

23. The method of claim 20, wherein the DNA encoding EPO is a genomic DNA.

24. The method of claim 20, wherein the DNA encoding the EPO is selected from the group consisting of SY, JM, SH and HE  
10 described in Fig. 5.

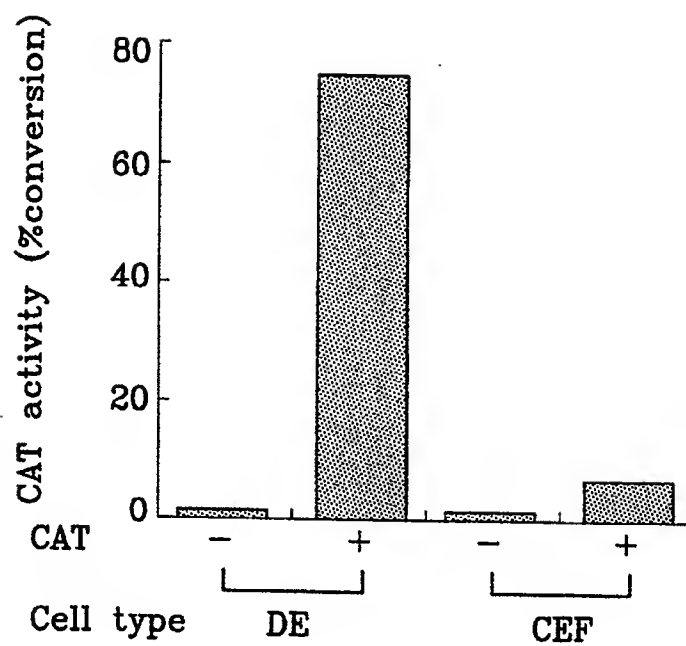
25. The method of claim 20, wherein the vector contains a promoter selected from the group consisting of SV40 early promoter, RSV LTR and HCMV MIEP.

26. An EPO genomic sequence selected from the group  
15 consisting of SY, JM, SH and HE described in Fig. 5.

27. An EPO amino acid sequence selected from the group consisting of JM, SH and HE described in Fig. 6.

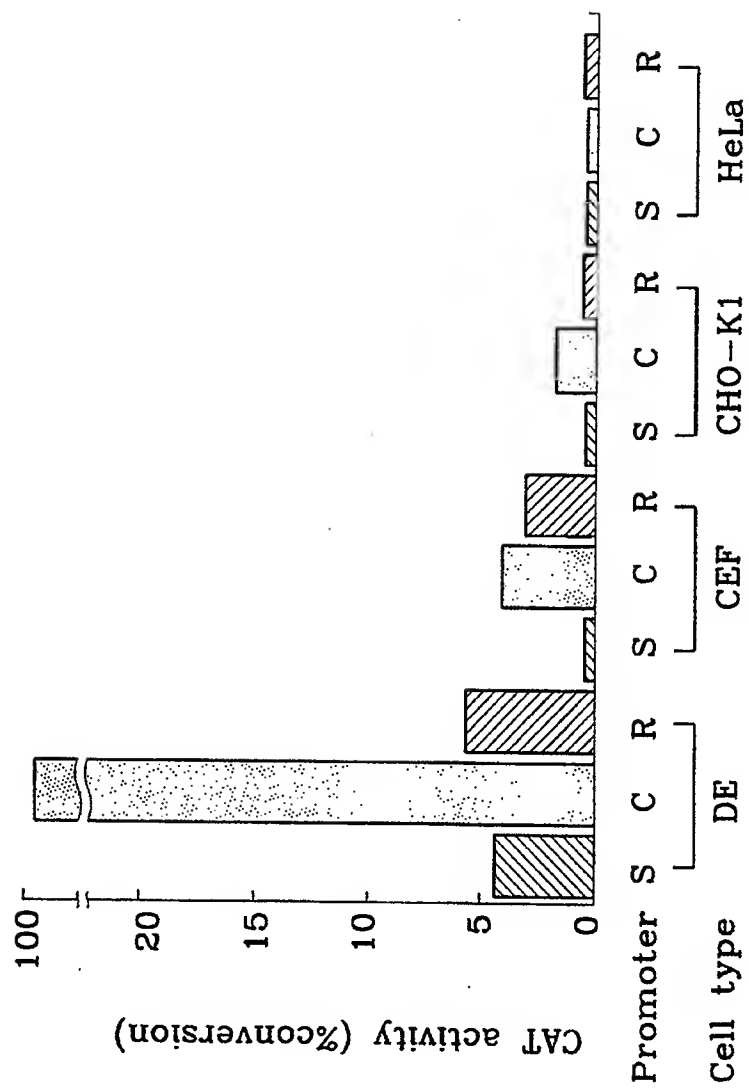
1/17

FIG. 1



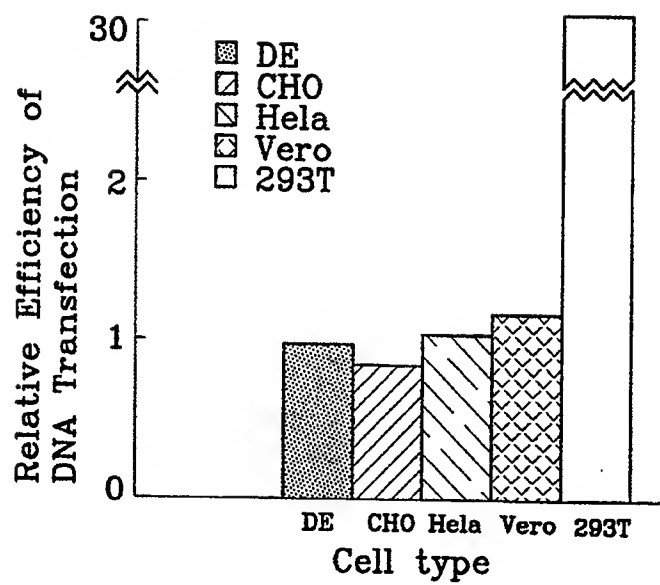
2/17

FIG.2



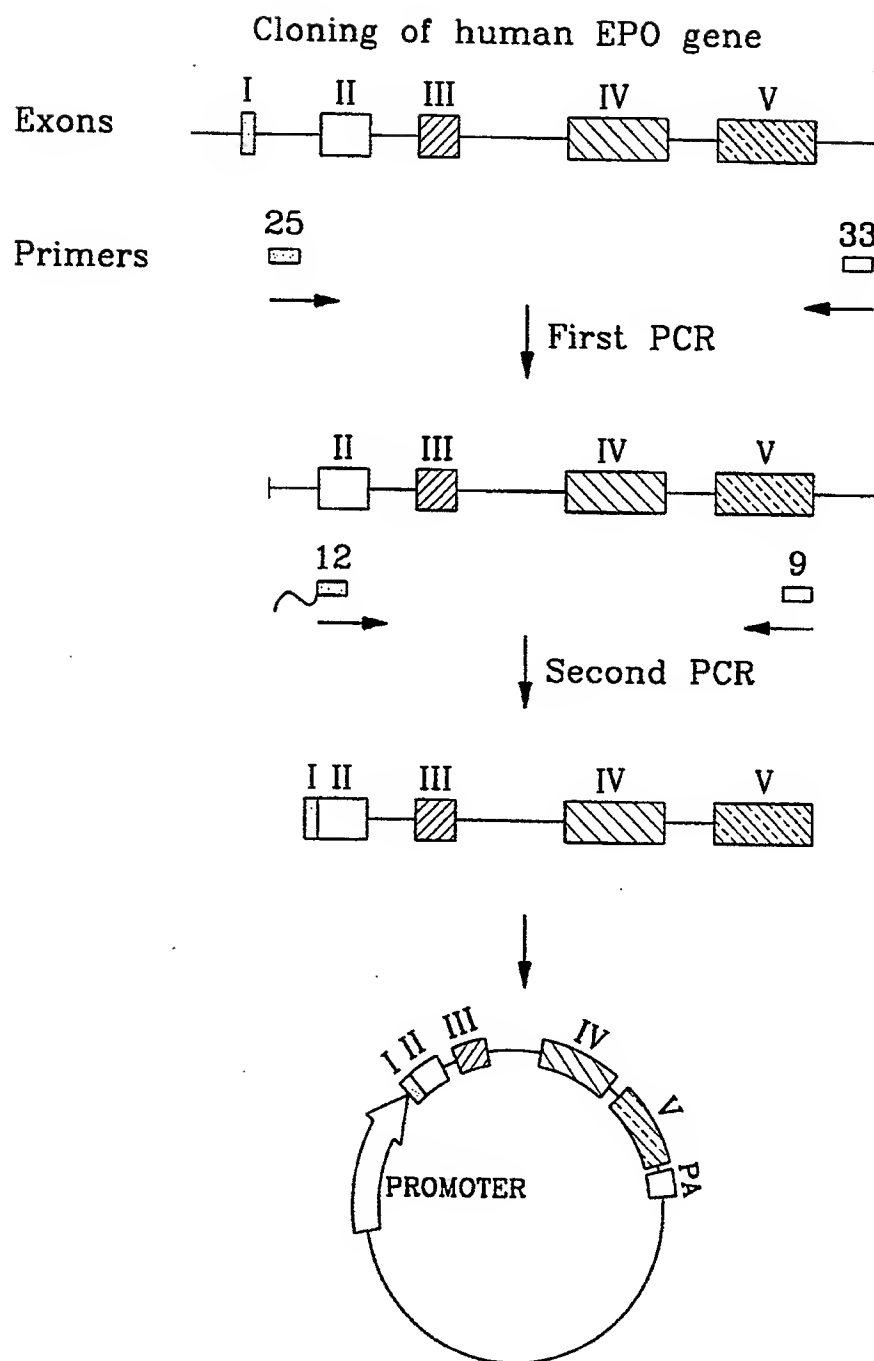
3/17

FIG.3



4/17

FIG.4



5/17

## FIG.5A

|       |                                                    |     |
|-------|----------------------------------------------------|-----|
| AM    | ATGGGGGTGCACGAATGTCCTGCCTGGCTGTGGCTTCTCCTGTCCCTGCT | 50  |
| GI    | ATGGGGGTGCACGAATGTCCTGCCTGGCTGTGGCTTCTCCTGTCCCTGCT | 50  |
| SY    | ATGGGGGTGCACGAATGTCCTGCCTGGCTGTGGCTTCTCCTGTCCCTGCT | 50  |
| JM    | ATGGGGGTGCACGAATGTCCTGCCTGGCTGTGGCTTCTCCTGTCCCTGCT | 50  |
| SH    | ATGGGGGTGCACGAATGTCCTGCCTGGCTGTGGCTTCTCCTGTCCCTGCT | 50  |
| HE    | ATGGGGGTGCACGAATGTCCTGCCTGGCTGTGGCTTCTCCTGTCCCTGCT | 50  |
| ***** |                                                    |     |
| AM    | GTCGCTCCCTCTGGGCCTCCAGTCCTGGGCGCCCCACCACGCCTCATCT  | 100 |
| GI    | GTCGCTCCCTCTGGGCCTCCAGTCCTGGGCGCCCCACCACGCCTCATCT  | 100 |
| SY    | GTCGCTCCCTCTGGGCCTCCAGTCCTGGGCGCCCCACCACGCCTCATCT  | 100 |
| JM    | GTCGCTCCCTCTGGGCCTCCAGTCCTGGGCGCCCCACCACGCCTCATCT  | 100 |
| SH    | GTCGCTCCCTCTGGGCCTCCAGTCCTGGGCGCCCCACCACGCCTCATCT  | 100 |
| HE    | GTCGCTCCCTCTGGGCCTCCAGTCCTGGGCGCCCCACCACGCCTCATCT  | 100 |
| ***** |                                                    |     |
| AM    | GTGACAGCCGAGTCCTGGAGAGGTACCTCTTGGAGGCCAAGGAGGCCGAG | 150 |
| GI    | GTGACAGCCGAGTCCTGGAGAGGTACCTCTTGGAGGCCAAGGAGGCCGAG | 150 |
| SY    | GTGACAGCCGAGTCCTGGAGAGGTACCTCTTGGAGGCCAAGGAGGCCGAG | 150 |
| JM    | GTGACAGCCGAGTCCTGGAGAGGTACCTCTTGGAGGCCAAGGAGGCCGAG | 150 |
| SH    | GTGACAGCCGAGTCCTGGAGAGGTACCTCTTGGAGGCCAAGGAGGCCGAG | 150 |
| HE    | GTGACAGCCGAGTCCTGGAGAGGTACCTCTTGGAGGCCAAGGAGGCCGAG | 150 |
| ***** |                                                    |     |
| AM    | AATATCACGGTGAGACCCCTTCCCAGCACATTCCACAGAACTCACGCTC  | 200 |
| GI    | AATATCACGGTGAGACCCCTTCCCAGCACATTCCACAGAACTCACGCTC  | 200 |
| SY    | AATATCACGGTGAGACCCCTTCCCAGCACATTCCACAGAACTCACGCTC  | 200 |
| JM    | AATATCACGGTGAGACCCCTTCCCAGCACATTCCACAGAACTCACGCTC  | 200 |
| SH    | AATATCACGGTGAGACCCCTTCCCAGCACATTCCACAGAACTCACGCTC  | 200 |
| HE    | AATATCACGGTGAGACCCCTTCCCAGCACATTCCACAGAACTCACGCTC  | 200 |
| ***** |                                                    |     |

6/17

## FIG.5B

|    |                                                    |     |
|----|----------------------------------------------------|-----|
| AM | AGGGCTTCAGGG-AACTCCTCCCAG-ATCCAGGAACCTGGCACTTGGTTT | 248 |
| GI | AGGGCTTCAGGG-AACTCCTCCCAG-ATCCAGGAACCTGGCACTTGGTTT | 248 |
| SY | AGGGCTTCAGGG-AACTCCTCCCAG-ATCCAGGAACCTGGCACTTGGTTT | 248 |
| JM | AGGGCTTCAGGG-AACTCCTCCCAG-ATCCAGGAACCTGGCACTTGGTTT | 248 |
| SH | AGGGCTTCAGGG-AACTCCTCCCAG-ATCCAGGAACCTGGCACTTGGTTT | 248 |
| HE | AGGGCTTCAGGG-AACTCCTCCCAG-ATCCAGGAACCTGGCACTTGGTTT | 250 |
|    | *****                                              |     |
| AM | GGGGTGGAGTTGGGAAGCTAGACACTGCCCCCTACATAAGAATAAGTCT  | 298 |
| GI | GGGGTGGAGTTGGGAAGCTAGACACTGCCCCCTACATAAGAATAAGTCT  | 298 |
| SY | GGGGTGGAGTTGGGAAGCTAGACACTGCCCCCTACATAAGAATAAGTCT  | 298 |
| JM | GGGGTGGAGTTGGGAAGCTAGACACTGCCCCCTACATAAGAATAAGTCT  | 298 |
| SH | GGGGTGGAGTTGGGAAGCTAGACACTGCCCCCTACATAAGAATAAGTCT  | 298 |
| HE | GGGGTGGAGTTGGGAAGCTAGACACTGCCCCCTACATAAGAATAAGTCT  | 300 |
|    | *****                                              |     |
| AM | GGTGGCCCCAAACCATACCTGGAACTAGGCAAGGAGCAAAGCCAGCAGA  | 348 |
| GI | GGTGGCCCCAAACCATACCTGGAACTAGGCAAGGAGCAAAGCCAGCAGA  | 348 |
| SY | GGTGGCCCCAAACCATACCTGGAACTAGGCAAGGAGCAAAGCCAGCAGA  | 348 |
| JM | GGTGGCCCCAAACCATACCTGGAACTAGGCAAGGAGCAAAGCCAGCAGA  | 348 |
| SH | GGTGGCCCCAAACCATACCTGGAACTAGGCAAGGAGCAAAGCCAGCAGA  | 348 |
| HE | GGTGGCCCCAAACCATACCTGGAACTAGGCAAGGAGCAAAGCCAGCAGA  | 350 |
|    | *****                                              |     |
| AM | TCCTACGGCCTGTGGCCAGGGCCAG-AGCCTTCAGGGACCCTTGACTCC  | 397 |
| GI | TCCTACGGCCTGTGGCCAGGGCCAG-AGCCTTCAGGGACCCTTGACTCC  | 395 |
| SY | TCCTACGGCCTGTGGCCAGGGCCAA-AGCCTTCAGGGACCCTTGACTCC  | 397 |
| JM | TCCTACGGCCTGTGGCCAGGGCCAGGAGCCTTCAGGGACCCTTGACTCC  | 398 |
| SH | TCCTACGGCCTGTGGCCAGGGCCA-AGCCTTCAGGGACCCTTGACTCC   | 397 |
| HE | TCCTACGGCCTGTGGCCAGGGCCA-AGCCTTCAGGGACCCTTGACTCC   | 399 |
|    | *****                                              |     |

7/17

## FIG.5C

|    |                                                     |     |
|----|-----------------------------------------------------|-----|
| AM | CCGGGCTGTGTGCATTTCAGACGGGCTGTGCTGAACACTGCAGCTTGAAT  | 447 |
| GI | CCGGGCTGTGTGCATTTCAGACGGGCTGTGCTGAACACTGCAGCTTGAAT  | 445 |
| SY | CCGGGCTGTGTGCATTTCAGACGGGCTGTGCTGAACACTGCAGCTTGAAT  | 447 |
| JM | CCGGGCTGTGTGCATTTCAGACGGGCTGTGCTGAACACTGCAGCTTGAAT  | 448 |
| SH | CCGGGCTGTGTGCATTTCAGACGGGCTGTGCTGAACACTGCAGCTTGAAT  | 447 |
| HE | CCGGGCTGTGTGCATTTCAGACGGGCTGTGCTGAACACTGCAGCTTGAAT  | 449 |
|    | *****                                               |     |
| AM | GAGAATATCACTGTCCAGACACCAAAGTTAATTTCTATGCCTGGAAGAG   | 497 |
| GI | GAGAATATCACTGTCCAGACACCAAAGTTAATTTCTATGCCTGGAAGAG   | 495 |
| SY | GAGAATATCACTGTCCAGACACCAAAGTTAATTTCTATGCCTGGAAGAG   | 497 |
| JM | GAGAATATCACTGTCCAGACACCAAAGTTAATTTCTATGCCTGGAAGAG   | 498 |
| SH | GAGAATATCACTGTCCAGACACCAAAGTTAATTTCTATGCCTGGAAGAG   | 497 |
| HE | GAGAATATCACTGTCCAGACACCAAAGTTAATTTCTATGCCTGGAAGAG   | 499 |
|    | **                                                  |     |
| AM | GATGGAGGTGAGTTCCTTTTTTTTTTTTTCCTTTCTTTTGGAGAATCT    | 547 |
| GI | GATGGAGGTGAGTTCCTTTTTTTTTTTTTCCTTTCTTTTGGAGAATCT    | 545 |
| SY | GATGGAGGTGAGTTCCTTTTTTTTTTTTTCCTTTCTTTTGGAGAATCT    | 547 |
| JM | GATGGAGGTGAGTTCCTTTTTTTTTTTTTCCTTTCTTTTGGAGAATCT    | 548 |
| SH | GATGGAGGTGAGTTCCTTTTTTTTTTTTTCCTTTCTTTTGGAGAATCT    | 545 |
| HE | GATGGAGGTGAGTTCCTTTTTTTTTTTTTCCTTTCTTTTGGAGAATCT    | 549 |
|    | *****                                               |     |
| AM | CATTTGCGAGCCTGATTTTGGATGAAAGGGAGAAATGATCGAGGGAAAGGT | 597 |
| GI | CATTTGCGAGCCTGATTTTGGATGAAAGGGAGAAATGATCGAGGGAAAGGT | 595 |
| SY | CATTTGCGAGCCTGATTTTGGATGAAAGGGAGAAATGATCGAGGGAAAGGT | 597 |
| JM | CATTTGCGAGCCTGATTTTGGATGAAAGGGAGAGTGATCGAGGGAAAGGT  | 598 |
| SH | CATTTGCGAGCCTGATTTTGGATGAAAGGGAGAAATGATCGAGGGAAAGGT | 595 |
| HE | CATTTGCGAGCCTGATTTTGGATGAAAGGGAGAAATGATCGAGGGAAAGGT | 599 |
|    | *****                                               |     |

8/17

## FIG.5D

|    |                                                    |     |
|----|----------------------------------------------------|-----|
| AM | AAAATGGAGCAGCAGAGATGAGGCTGCCTGGGCGCAGAGGCTCACGTCTA | 647 |
| GI | AAAATGGAGCAGCAGAGATGAGGCTGCCTGGGCGCAGAGGCTCACGTCTA | 645 |
| SY | AAAATGGAGCAGCAGAGATGAGGCTGCCTGGGCGCAGAGGCTCACGTCTA | 647 |
| JM | AAAATGGAGCAGCAGAGATGAGGCTGCCTGGGCGCAGAGGCTCACGTCTA | 648 |
| SH | AAAATGGAGCAGCAGAGATGAGGCTGCCTGGGCGCAGAGGCTCACGTCTA | 645 |
| HE | AAAATGGAGCAGCAGAGATGAGGCTGCCTGGGCGCAGAGGCTCACGTCTA | 649 |
|    | *****                                              |     |
| AM | TAATCCCAGGCTGAGATGGCCGAGATGGGAGAATTGCTTGAGCCCTGGAG | 697 |
| GI | TAATCCCAGGCTGAGATGGCCGAGATGGGAGAATTGCTTGAGCCCTGGAG | 695 |
| SY | TAATCCCAGGCTGAGATGGCCGAGATGGGAGAATTGCTTGAGCCCTGGAG | 697 |
| JM | TAATCCCAGGCTGAGATGGCCGAGATGGGAGAATTGCTTGAGCCCTGGAG | 698 |
| SH | TAATCCCAGGCTGAGATGGCCGAGATGGGAGAATTGCTTGAGCCCTGGAG | 695 |
| HE | TAATCCCAGGCTGAGATGGCCGAGATGGGAGAATTGCTTGAGCCCTGGAG | 699 |
|    | *****                                              |     |
| AM | G TTCAGACCAACCTAGGCAGCATAGTGAGATCCCCATCTCTACAAACAT | 747 |
| GI | G TTCAGACCAACCTAGGCAGCATAGTGAGATCCCCATCTCTACAAACAT | 747 |
| SY | G TTCAGACCAACCTAGGCAGCATAGTGAGATCCCCATCTCTACAAACAT | 747 |
| JM | G TTCAGACCAACCTAGGCAGCATAGTGAGATCCCCATCTCTACAAACAT | 748 |
| SH | G TTCAGACCAACCTAGGCAGCATAGTGAGATCCCCATCTCTACAAACAT | 745 |
| HE | G TTCAGACCAACCTAGGCAGCATAGTGAGATCCCCATCTCTACAAACAT | 749 |
|    | *****                                              |     |
| AM | TTAAAAAAATTAGTCAGGTGAAGTGGTGCATGGTGGTAGTCCCAGATATT | 797 |
| GI | TTAAAAAAATTAGTCAGGTGAAGTGGTGCATGGTGGTAGTCCCAGATATT | 795 |
| SY | TTAAAAAAATTAGTCAGGTGAAGTGGTGCATGGTGGTAGTCCCAGATATT | 797 |
| JM | TTAAAAAAATTAGTCAGGTGAAGTGGTGCATGGTGGTAGTCCCAGATATT | 798 |
| SH | TTAAAAAAATTAGTCAGGTGAAGTGGTGCATGGTGGTAGTCCCAGATATT | 795 |
| HE | TTAAAAAAATTAGTCAGGTGAAGTGGTGCATGGTGGTAGTCCCAGATATT | 799 |
|    | *****                                              |     |

9/17

## FIG.5E

|    |                                                      |     |
|----|------------------------------------------------------|-----|
| AM | TGGAAGGCTGAGGCGGGAGGATCGCTTGAGCCCAGGAATTTGAGGCTGCA   | 847 |
| GI | TGGAAGGCTGAGGCGGGAGGATCGCTTGAGCCCAGGAATTTGAGGCTGCA   | 845 |
| SY | TGGATGGCTGAGGCGGGAGGATCGCTTGAGCCCAGGAATTTGAGGCTGCA   | 847 |
| JM | TGGAAGGCTGAGGCGGGAGGATCGCTTGAGCCCAGGAATTTGAGGCTGCG   | 848 |
| SH | TGGAAGGCTGAGGCGGGAGGATCGCTTGAGCCCAGGAATTTGGGGCTGCA   | 845 |
| HE | TGGAAGGCTGAGGCGGGAGGATCGCTTGAGCCCAGGAATTTGAGGCTGCA   | 849 |
|    | *****                                                |     |
| AM | GTGAGCTGTGATCACACCACTGCACTCCAGCCTCAGTGACAGAGTGAGGC   | 897 |
| GI | GTGAGCTGTGATCACACCACTGCACTCCAGCCTCAGTGACAGAGTGAGGC   | 895 |
| SY | GTGAGCTGTGATCACACCACTGCACTCCAGCCTCAGTGACAGAGTGAGGC   | 897 |
| JM | GTGAGCTGTGATCACACCACTGCACTCCAGCCTCAGTGACAGAGTGAGGC   | 898 |
| SH | GTGAGCTGTGATCACACCACTGCACTCCAGCCTCAGTGACAGAGTGAGGC   | 895 |
| HE | GTGAGCTGTGATCACACCACTGCACTCCAGCCTCAGTGACAGAGTGAGGC   | 899 |
|    | *****                                                |     |
| AM | CCTGTCTCAAAAAAGAAAAGAAAAAGAAAAATATGAGGGCTGTATGGA     | 947 |
| GI | CCTGTCTCAAAAAAGAAAAGAAAAAGAAAAATATGAGGGCTGTATGGA     | 945 |
| SY | CCTGTCTCAAAAAAGAAAAGAAAAAGAAAAATATGAGGGCTGTATGGA     | 947 |
| JM | CCTGTCTCAAAAAAGAAAAGAAAAAGAAAAATATGAGGGCTGTATGGA     | 948 |
| SH | CCTGTCTCAAAAAAGAAAAGAAAAAGAAAAATATGAGGGCTGTATGGA     | 945 |
| HE | CCTGTCTCAAAAAAGAAAAGAAAAAGAAAAATATGAGGGCTGTATGGA     | 949 |
|    | *****                                                |     |
| AM | ATACATTCATTATTCATTCACCTCACTCACTCACTCATTCATTCAATTCATT | 997 |
| GI | ATACATTCATTATTCATTCACCTCACTCACTCACTCATTCATTCAATTCATT | 995 |
| SY | ATACATTCATTATTCATTCACCTCACTCACTCACTCATTCATTCAATTCATT | 997 |
| JM | ATACATTCATTATTCATTCACCTCACTCACTCACTCATTCATTCAATTCATT | 998 |
| SH | ATACATTCATTATTCATTCACCTCACTCACTCACTCATTCATTCAATTCATT | 995 |
| HE | ATACATTCATTATTCATTCACCTCACTCACTCACTCATTCATTCAATTCATT | 999 |
|    | *****                                                |     |

10/17

## FIG.5F

|         |                                                    |      |
|---------|----------------------------------------------------|------|
| AM      | CATTCAACAAGTCTTATTGCATACCTTCTGTTTGCTCAGCTTGGTGCTTG | 1047 |
| GI      | CATTCAACATGTCTTATTGCATACCTTCTGTTTGCTCAGCTTGGTGCTTG | 1045 |
| SY      | CATTCAACAAGTCTTATTGCATACCTTCTGTTTGCTCAGCTTGGTGCTTG | 1047 |
| JM      | CATTCAACAAGTCTTATTGCATACCTTCTGTTTGCTCAGCTTGGTGCTTG | 1048 |
| SH      | CATTCAACAAGTCTTATTGCATACCTTCTGTTTGCTCAGCTTGGTGCTTG | 1045 |
| HE      | CATTCAACAAGTCTTATTGCATACCTTCTGTTTGCTCAGCTTGGTGCTTG | 1049 |
| *****   |                                                    |      |
| AM      | GGGCTGCTGAGGGGCAGGAGGGAGAGGGTGACATGGGTCACTGACTCCC  | 1097 |
| GI      | GGGCTGCTGAGGGGCAGGAGGGAGAGGGTGACATCCCTCAGCTGACTCCC | 1095 |
| SY      | GGGCTGCTGAGGGGCAGGAGGGAGAGGGTGACATGGGTCACTGACTCCC  | 1097 |
| JM      | GGGCTGCTGAGGGGCAGGAGGGAGAGGGTGACATGGGTCACTGACTCCC  | 1098 |
| SH      | GGGCTGCTGAGGGGCAGGAGGGAGAGGGTGACATGGGTCACTGACTCCC  | 1095 |
| HE      | GGGCTGCTGAGGGGCAGGAGGGAGAGGGTGACATGGGTCACTGACTCCC  | 1099 |
| ** ** * |                                                    |      |
| AM      | AGAGTCCACTCCCTGTAGGTCGGGCAGCAGGCCGTAGAAGTCTGGCAGGG | 1147 |
| GI      | AGAGTCCACTCCCTGTAGGTCGGGCAGCAGGCCGTAGAAGTCTGGCAGGG | 1145 |
| SY      | AGAGTCCACTCCCTGTAGGTCGGGCAGCAGGCCGTAGAAGTCTGGCAGGG | 1147 |
| JM      | AGAGTCCACTCCCTGTAGGTCGGGCAGCAGGCCGTAGAAGTCTGGCAGGG | 1148 |
| SH      | AGAGTCCACTCCCTGTAGGTCGGGCAGCAGGCCGTAGAAGTCTGGCAGGG | 1145 |
| HE      | AGAGTCCACTCCCTGTAGGTCGGGCAGCAGGCCGTAGAAGTCTGGCAGGG | 1149 |
| *****   |                                                    |      |
| AM      | CCTGGCCCTGCTGTCGGAACTGTCTGCGGGGCCAGGCCCTGTTGGTCA   | 1197 |
| GI      | CCTGGCCCTGCTGTCGGAACTGTCTGCGGGGCCAGGCCCTGTTGGTCA   | 1195 |
| SY      | CCTGGCCCTGCTGTCGGAACTGTCTGCGGGGCCAGGCCCTGTTGGTCA   | 1197 |
| JM      | CCTGGCCCTGCTGTCGGAACTGTCTGCGGGGCCAGGCCCTGTTGGTCA   | 1198 |
| SH      | CCTGGCCCTGCTGTCGGAACTGTCTGCGGGGCCAGGCCCTGTTGGTCA   | 1195 |
| HE      | CCTGGCCCTGCTGTCGGAACTGTCTGCGGGGCCAGGCCCTGTTGGTCA   | 1199 |
| *****   |                                                    |      |

11/17

## FIG.5G

|    |                                                     |      |
|----|-----------------------------------------------------|------|
| AM | ACTCTTCCCAGCCGTGGGAGCCCCCTGCAGCTGCATGTGGATAAAGCCGTC | 1247 |
| GI | ACTCTTCCCAGCCGTGGGAGCCCCCTGCAGCTGCATGTGGATAAAGCCGTC | 1245 |
| SY | ACTCTTCCCAGCCGTGGGAGCCCCCTGCAGCTGCATGTGGATAAAGCCGTC | 1247 |
| JM | ACTCTTCCCAGCCGTGGGAGCCCCCTGCAGCTGCATGTGGATAAAGCCGTC | 1248 |
| SH | ACTCTTCCCAGCCGTGGGAGCCCCCTGCAGCTGCATGTGGATAAAGCCGTC | 1245 |
| HE | ACTCTTCCCAGCCGTGGGAGCCCCCTGCAGCTGCATGTGGATAAAGCCGTC | 1249 |
|    | *****                                               |      |
| AM | AGTGGCCTTCGCAGCCTCACCACCTCTGCTTCGGGCTCTGGGAGCCAGGT  | 1297 |
| GI | AGTGGCCTTCGCAGCCTCACCACCTCTGCTTCGGGCTCTGGGAGCCAGGT  | 1295 |
| SY | AGTGGCCTTCGCAGCCTCACCACCTCTGCTTCGGGCTCTGGGAGCCAGGT  | 1297 |
| JM | AGTGGCCTTCGCAGCCTCACCACCTCTGCTTCGGGCTCTGGGAGCCAGGT  | 1298 |
| SH | AGTGGCCTTCGCAGCCTCACCACCTCTGCTTCGGGCTCTGGGAGCCAGGT  | 1295 |
| HE | AGTGGCCTTCGCAGCCTCACCACCTCTGCTTCGGGCTCTGGGAGCCAGGT  | 1299 |
|    | *****                                               |      |
| AM | GAGTAGGAGCGGACACTTCTGCTTGCCCTTTCTGTAAGAAGGGGAGAAGG  | 1347 |
| GI | GAGTAGGAGCGGACACTTCTGCTTGCCCTTTCTGTAAGAAGGGGAGAAGG  | 1345 |
| SY | GAGTAGGAGCGGACACTTCTGCTTGCCCTTTCTGTAAGAAGGGGAGAAGG  | 1347 |
| JM | GAGTAGGAGCGGACACTTCTGCTTGCCCTTTCTGTAAGAAGGGGAGAAGG  | 1348 |
| SH | GAGTAGGAGCGGACACTTCTGCTTGCCCTTTCTGTAAGAAGGGGAGAAGG  | 1345 |
| HE | GAGTAGGAGCGGACACTTCTGCTTGCCCTTTCTGTAAGAAGGGGAGAAGG  | 1349 |
|    | *****                                               |      |
| AM | GTCTTGCTAAGGAGTACAGGAACGTCCGTATTCCTTCCCTTTCTGTGGC   | 1397 |
| GI | GTCTTGCTAAGGAGTACAGGAACGTCCGTATTCCTTCCCTTTCTGTGGC   | 1395 |
| SY | GTCTTGCTAAGGAGTACAGGAACGTCCGTATTCCTTCCCTTTCTGTGGC   | 1397 |
| JM | GTCTTGCTAAGGAGTACAGGAACGTCCGTATTCCTTCCCTTTCTGTGGC   | 1398 |
| SH | GTCTTGCTAAGGAGTACAGGAACGTCCGTATTCCTTCCCTTTCTGTGGC   | 1395 |
| HE | GTCTTGCTAAGGAGTACAGGAACGTCCGTATTCCTTCCCTTTCTGTGGC   | 1399 |
|    | *****                                               |      |

12/17

## FIG.5H

|       |                                                   |      |
|-------|---------------------------------------------------|------|
| AM    | ACTGCAGCGACCTCCTGTTTTCTCCTTGGCAGAAGGAAGCCATCTCCCT | 1447 |
| GI    | ACTGCAGCGACCTCCTGTTTTCTCCTTGGCAGAAGGAAGCCATCTCCCT | 1445 |
| SY    | ACTGCAGCGACCTCCTGTTTTCTCCTTGGCAGAAGGAAGCCATCTCCCT | 1447 |
| JM    | ACTGCAGCGACCTCCTGTTTTCTCCTTGGCAGAAGGAAGCCATCTCCCT | 1448 |
| SH    | ACTGCAGCGACCTCCTGTTTTCTCCTTGGCAGAAGGAAGCCATCTCCCT | 1445 |
| HE    | ACTGCAGCGACCTCCTGTTTTCTCCTTGGCAGAAGGAAGCCATCTCCCT | 1449 |
| ***** |                                                   |      |
| AM    | CCAGATGCGGCCTCAGCTGCTCCACTCCGAACAATCACTGCTGACTTT  | 1497 |
| GI    | CCAGATGCGGCCTCAGCTGCTCCACTCCGAACAATCACTGCTGACTTT  | 1495 |
| SY    | CCAGATGCGGCCTCAGCTGCTCCACTCCGAACAATCACTGCTGACTTT  | 1497 |
| JM    | CCAGATGCGGCCTCAGCTGCTCCACTCCGAACAATCACTGCTGACTTT  | 1498 |
| SH    | CCAGATGCGGCCTCAGCTGCTCCACTCCGAACAATCACTGCTGACTTT  | 1495 |
| HE    | CCAGATGCGGCCTCAGCTGCTCCACTCCGAACAATCACTGCTGACTTT  | 1499 |
| ***** |                                                   |      |
| AM    | CCGCAAACCTCTCCGAGTCTACTCCAATTTCTCCGGGGAGAGCTGAAGC | 1547 |
| GI    | CCGCAAACCTCTCCGAGTCTACTCCAATTTCTCCGGGGAGAGCTGAAGC | 1545 |
| SY    | CCGCAAACCTCTCCGAGTCTACTCCAATTTCTCCGGGGAGAGCTGAAGC | 1547 |
| JM    | CCGCAAACCTCTCCGAGTCTACTCCAATTTCTCCGGGGAGAGCTGAAGC | 1548 |
| SH    | CCGCAAACCTCTCCGAGTCTACTCCAATTTCTCCGGGGAGAGCTGAAGC | 1545 |
| HE    | CCGCAAACCTCTCCGAGTCTACTCCAATTTCTCCGGGGAGAGCTGAAGC | 1549 |
| ***** |                                                   |      |
| AM    | TGTACACAGGGGAGGCCTGCAGGACAGGGGACAGATGA            | 1584 |
| GI    | TGTACACAGGGGAGGCCTGCAGGACAGGGGACAGATGA            | 1582 |
| SY    | TGTACACAGGGGAGGCCTGCAGGACAGGGGACAGATGA            | 1585 |
| JM    | TGTACACAGGGGAGGCCTGCAGGACAGGGGACAGATGA            | 1585 |
| SH    | TGTACACAGGGGAGGCCTGCAGGACAGGGGACAGATGA            | 1583 |
| HE    | TGTACACAGGGGAGGCCTGCAGGACAGGGGACAGATGA            | 1586 |
| ***** |                                                   |      |

13/17

## FIG.6

|       |                                                    |    |
|-------|----------------------------------------------------|----|
| AM/GI | MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAE | 50 |
| SY    | MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAE | 50 |
| JM    | MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAE | 50 |
| SH    | MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDRVLERYLLEAKEAE  | 50 |
| HE    | MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAE | 50 |
| ***** |                                                    |    |

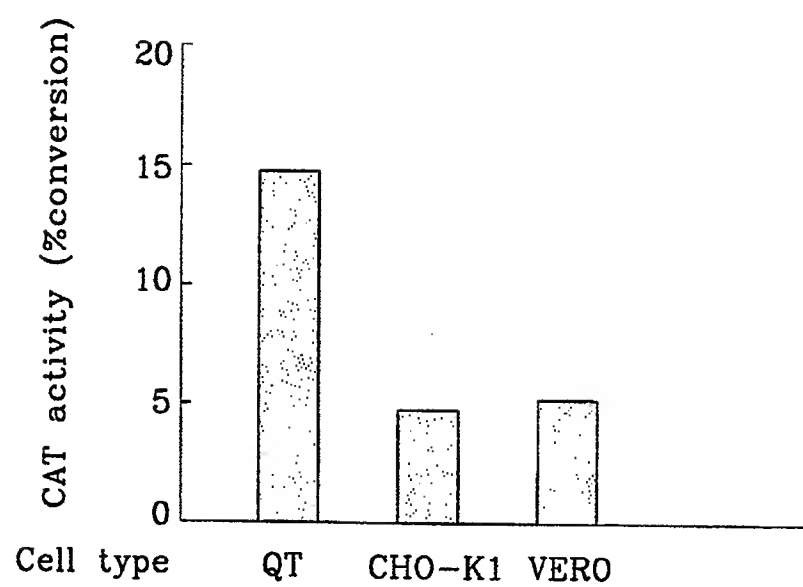
|           |                                                     |     |
|-----------|-----------------------------------------------------|-----|
| AM/GI     | NITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQAQVEVWQGLALLSEA | 100 |
| SY        | NITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQAQVEVWQGLALLSEA | 100 |
| JM        | NITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQAQVEVWQGLALLSEA | 100 |
| SH        | NITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQAQVEVWQGLALLSES | 100 |
| HE        | NITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQAQVEVWQGLALLSEA | 100 |
| *** ***** |                                                     |     |

|       |                                                   |     |
|-------|---------------------------------------------------|-----|
| AM/GI | VLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTLLRALGAQKEAISPPD | 150 |
| SY    | VLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTLLRALGAQKEAISPPD | 150 |
| JM    | VLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTLLRALGAQKEAISPPD | 150 |
| SH    | VLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTLLRALGAQKEAISPPD | 150 |
| HE    | VLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTLLRALGAQKEAISPPD | 150 |
| ***** |                                                   |     |

|       |                                             |     |
|-------|---------------------------------------------|-----|
| AM/GI | AASAAPLRTITADTFRKLFRVYSNFLRGKIKLYTGEACRTGDR | 193 |
| SY    | AASAAPLRTITADTFRKLFRVYSNFLRGKIKLYTGEACRTGDR | 193 |
| JM    | AASAAPLRTITADTFRKLFRVYSNFLRGKIKLYTGEACRTGDR | 193 |
| SH    | AASAAPLRTITADTFRKLFRVYSNFLRGKIKLYTGEACRTGDR | 193 |
| HE    | AASAAPLRTITADTFRKLFRVYSNFLRGKIKLYTGEACRTGDR | 193 |
| ***** |                                             |     |

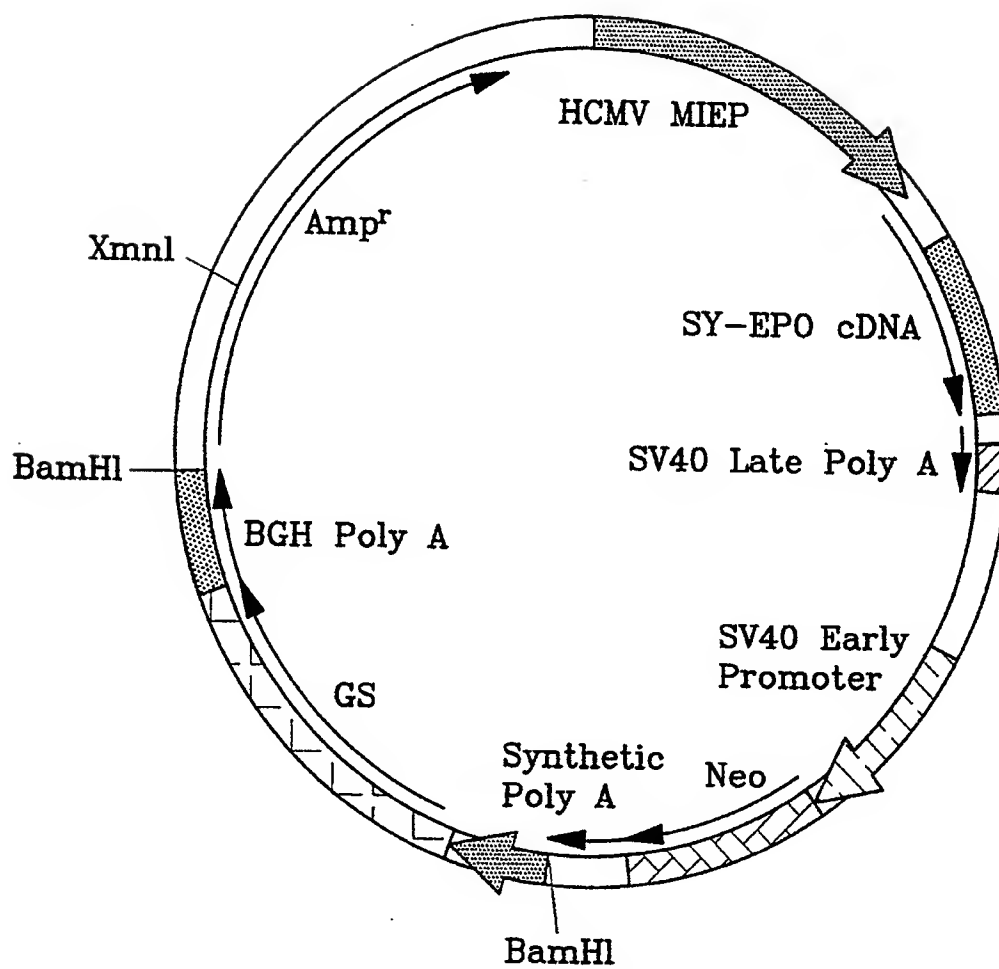
14/17

FIG. 7



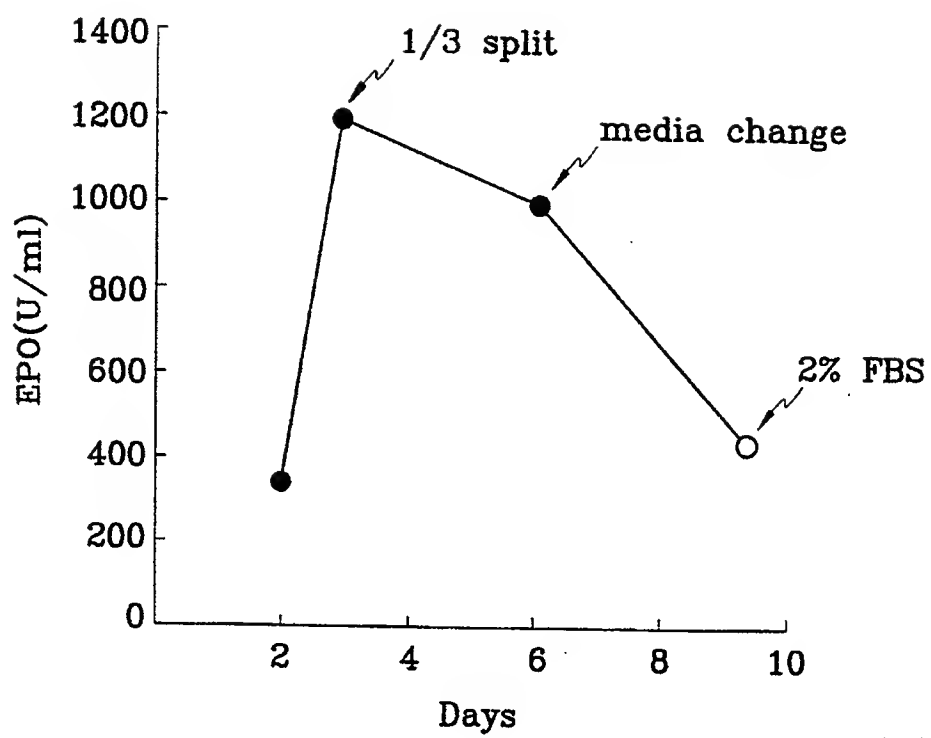
15/17

FIG.8



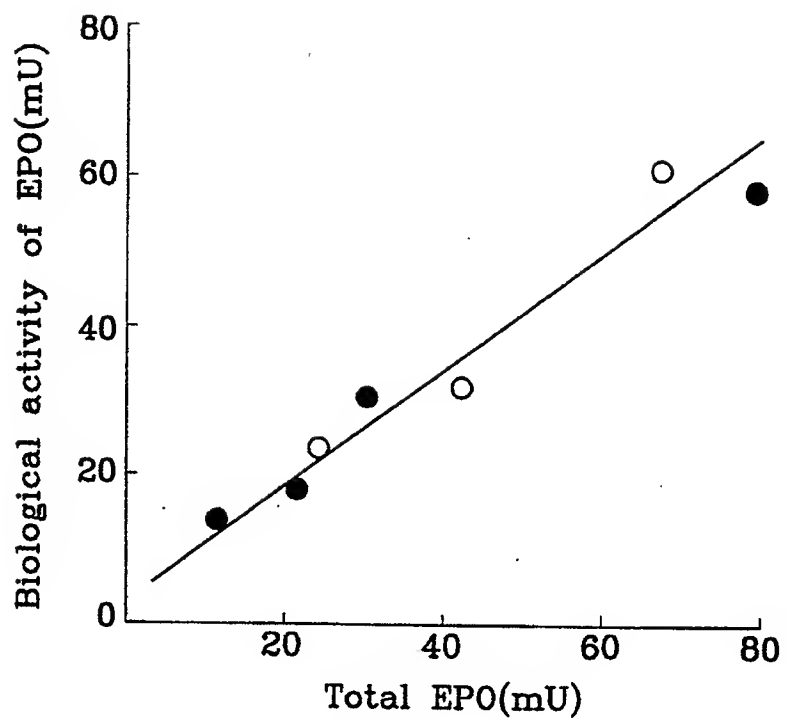
16/17

FIG.9



17/17

FIG.10



## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/KR 96/00145

## A. CLASSIFICATION OF SUBJECT MATTER

IPC<sup>6</sup>: C 12 N 15/11, 15/12

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC<sup>6</sup>: C 12 N 15/11, 15/12

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched  
1e

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

WPIL

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages                                               | Relevant to claim No. |
|-----------|----------------------------------------------------------------------------------------------------------------------------------|-----------------------|
| X         | WO 89/ 03 877 A1 (INSTITUT INTERNATIONAL DE LA RECHERCHE AGRONOMIQUE (INRA)) 05 May 1989 (05.05.89), claims 1,18,21,23.<br>----- | 1,8                   |

☐ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

\* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date  
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search  
08 November 1996 (08.11.96)

Date of mailing of the international search report  
05 December 1996 (05.12.96)

Name and mailing address of the ISA/ AT  
AUSTRIAN PATENT OFFICE  
Kohlmarkt 8-10  
A-1014 Vienna  
Facsimile No. 1/53424/535

Authorized officer  
Wolf  
Telephone No. 1/53424/133

**INTERNATIONAL SEARCH REPORT**  
Information on patent family members

International application No.

PCT/KR 96/00145

| la Recherchebericht<br>angeführtes Patentdokument<br>Patent document cited<br>in search report<br>Document de brevet cité<br>dans le rapport de recherche | Datum der<br>Veröffentlichung<br>Publication<br>date<br>Date de<br>publication | Mitglied(er) der<br>Patentfamilie<br>Patent family<br>member(s)<br>Membre(s) de la<br>famille de brevets | Datum der<br>Veröffentlichung<br>Publication<br>date<br>Date de<br>publication |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------|
| WO A1 8903877                                                                                                                                             | 05-05-89                                                                       | EP A1 395658                                                                                             | 07-11-90                                                                       |
|                                                                                                                                                           |                                                                                | FR A1 2622207                                                                                            | 26-04-89                                                                       |
|                                                                                                                                                           |                                                                                | FR B1 2622207                                                                                            | 09-03-90                                                                       |
|                                                                                                                                                           |                                                                                | JP T2 503839                                                                                             | 29-08-91                                                                       |
|                                                                                                                                                           |                                                                                | US A 5252465                                                                                             | 12-10-93                                                                       |

se